AFRL-HE-WP-TP-2006-0095

AIR FORCE RESEARCH LABORATORY

Performance and Psychophysiological Measures of Fatigue Effects on Aviation Related Tasks of Varying Difficulty

Glenn F. Wilson
Christopher Russell

Human Effectiveness Directorate
Wright-Patterson AFB OH

John Caldwell

Human Effectiveness Directorate
Brooks City Base TX

January 2006

Interim Report for August 2005 to December 2005

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright-Patterson AFB, OH 45433-7022

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
January 2006	Interim	August 2005 to December 2005

4. TITLE AND SUBTITLE
Performance and Psychophysiological Measures of Fatigue Effects on Aviation Related Tasks of Varying Difficulty

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
George Wilson *
Christopher Russell *
John Caldwell **

5d. PROJECT NUMBER
7184

5e. TASK NUMBER
08

5f. WORK UNIT NUMBER
64

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
*Human Effectiveness Directorate, WPAFB OH
** Human Effectiveness Directorate, Brooks City Base TX

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright Patterson AFB OH 45433-7022

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/HECP

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-HE-WP-TP-2006-0095

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited. Cleared by AFRL/WS-05-2774 on 13 December 2005.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Fatigue is a well known stressor in aviation operations and its interaction with mental workload needs to be understood. Performance, psychophysiological, and subjective measures were collected during performance of three tasks of increasing complexity. A psychomotor vigilance task, multi-attribute task battery and an uninhabited air vehicle task were performed five times during one night's sleep loss. EEG, ECG and pupil area were recorded during task performance. Performance decrements were found at the next to last and/or last testing session. The EEG showed concomitant changes. The degree of impairment was at least partially dependent on the task being performed and the performance variable assessed.

15. SUBJECT TERMS
Fatigue, psychophysiology, EEG, ECG, performance

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Glenn Wilson
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	SAR	52	**19b. TELEPONE NUMBER** *(Include area code)*

STANDARD FORM 298 Back (Rev. 8/98)

Running head: FATIGUE EFFECTS ON PERFORMANCE AND PSYCHOPHYSIOLOGY

Performance and Psychophysiological Measures of Fatigue Effects on Aviation Related Tasks of Varying Difficulty

Glenn F. Wilson

AFRL/HECP, Wright-Patterson AFB, OH

John A. Caldwell

AFRL/HEPF, Brooks AFB, TX

Christopher A. Russell

U.S. Air Force Research Laboratory

Abstract

Fatigue is a well known stressor in aviation operations and its interaction with mental workload needs to be understood. Performance, psychophysiological, and subjective measures were collected during performance of three tasks of increasing complexity. A psychomotor vigilance task, multi-attribute task battery and an uninhabited air vehicle task were performed five times during one night's sleep loss. EEG, ECG and pupil area were recorded during task performance. Performance decrements were found at the next to last and/or last testing session. The EEG showed concomitant changes. The degree of impairment was at least partially dependent on the task being performed and the performance variable assessed.

Keywords: Fatigue, psychophysiology, EEG, ECG, performance

INTRODUCTION

In a variety of real-world settings, fatigue stemming from long duty hours, insufficient sleep, and circadian factors can seriously degrade both the alertness and performance of operators (Akerstedt, 1995). In aviation-related occupations, on-the-job sleepiness is particularly dangerous because fatigue-related errors in the cockpit can lead to crew-member and passenger fatalities as well as the loss of the airframe itself (Caldwell and Caldwell, 2003). Furthermore, in military contexts, aviator fatigue can result in the failure to acquire, engage, and destroy enemy targets, or worse, to result in incorrect targeting and destruction of non-threatening ("friendly') assets in the air or on the ground.

Unfortunately, aircrew fatigue is a growing problem in flight-related operations due predominantly to crew scheduling and workload factors that curtail sleep and increase stress (Bourgeois-Bougrine et al., 2003; Rosekind et al., 2000; Co et al., 1999; Caldwell and Gilreath, 2001). Although air vehicles and other equipment can function nonstop for extended periods without adverse effects, human operators cannot work and think effectively in the absence of adequate restorative sleep (Horne, 1978). Sleep deprivation rapidly produces high levels of fatigue that can lead to dangerously compromised performance (Krueger, 1989). Deleterious effects include degradations in response accuracy and speed, the unconscious acceptance of lower standards of performance, impairments in the capacity to integrate information, and narrowing of attention (Perry, 1974). Performance becomes less consistent while overall vigilance deteriorates (Dinges, 1990), and operator sleepiness adversely

impacts the ability to retain new information and to detect unwanted system deviations (Falleti et al., 2003). In fact, it now appears that overly-tired aviators may face operational risks similar to those posed by alcohol intoxication. Recent studies have established that from 17-25 hours of sustained wakefulness can produce an array of performance deficits equivalent to those observed with blood alcohol concentrations (BAC's) of 0.05% to 0.10%--the legal limits for driver intoxication (Dawson and Reid, 1997; Lamond and Dawson, 1999).

Pilots flying at night or during the predawn hours are especially vulnerable to fatigue-related cognitive lapses, or even worse, "micro-sleeps"—brief periods during which sleep uncontrollably and often unconsciously intrudes into wakefulness. Moore-Ede (1993) found that aviators engaged in simulator flights during the predawn hours experienced a tenfold increase in the number of inadvertent sleep lapses, and during this same time (in which the micro-sleeps were most prominent), pilots made the greatest number of performance errors. Wright and McGown (2001) found that while sleepiness (both during the daytime and the night) was increasingly problematic as a function of increased flight duration, occurrences of outright electrophysiologically-measured sleep episodes were more frequent on flights with late-night departures than on flights that departed earlier in the day. These findings corroborated those of Rosekind et al. (1994) who determined that long-haul commercial airline pilots were particularly plagued by inadvertent (and often unrecognized) periods of dozing off as well as concurrent decrements in vigilance performance and subjective alertness ratings during night flights in comparison to flights during the daytime hours.

More recently, Caldwell et al. (2003) found that even well-trained Air Force fighter pilots are susceptible to the adverse effects of fatigue during extended periods of wakefulness. Psychological mood, self-reported alertness, central-nervous-system activation, and basic cognition were substantially impaired following 22 or more hours without sleep, and objectively-measured piloting skills were degraded by more than 40 percent due to aviator fatigue.

Thus, it is clear that "traditional" piloting skills (i.e., those involved in flying typical civilian and military aircraft) are significantly degraded by fatigue. It is equally clear that unless pilots and crew members are placed on well-planned work/rest schedules, there is a high probability that the crew, the airframe, and traditional air missions may be jeopardized by fatigue. However, at present, there is comparatively little understanding about the degree to which newly-developed aviation tasks might be compromised by the sleep loss and circadian disruptions associated with extended duty schedules. At present, the U.S. military is in the process of developing and fielding unmanned air vehicles (UAV's) that are controlled by remote, ground-based operators. Furthermore, the air vehicle piloting tasks themselves are sometimes far different than those required of more traditional pilots. For instance, some types of UAV operators may not be tasked with the responsibility for performing takeoffs and landings or for maintaining specific UAV flight parameters. Instead, the UAV may be automatically piloted by computers, while the operator is sifting through reconnaissance data, targeting weapons, monitoring various UAV systems, and completing a number of other tasks that have traditionally been handled by other aircrew members in conventional aircraft

operations. The effects of fatigue and circadian disruptions on UAV operations are not clear and should be investigated as the UAV presence increases in the inventory.

Furthermore, the interaction between cognitive workload and fatigue are not well understood. On the one hand, people subjectively associate high workload with greater fatigue (Akerstedt, 2004). On the other hand, some performance-based studies have not shown that fatigue-related difficulties are more severe on hard versus easy tasks (Wilkinson, 1964; Xue-liang, 2004). A study of air-traffic-control performance, while demonstrating the importance of time pressure and workload, offered only partial support for the assumption that the effects of sleep loss were dependent on the level of cognitive challenge posed by a task (Lichacz, 2005).

Because of the increasing use of unmanned vehicles the effects of fatigue on their operators must be studied. Further, the relationship between task performance and the functional state of the operator should be understood. Extensive research has shown that psychophysiological measures can be used to investigate operator functional states such as cognitive workload and fatigue (Kramer, 1991; Wilson & Eggemeier, 1991). Using nonintrusive psychophysiological measures it may be possible to design monitoring and alerting systems that will help maintain operator performance under fatigue conditions (Byrne & Parasuraman, 1996). The cognitive demands experienced by UAV operators vary over a wide range of task complexities. At one end of this spectrum are the less demanding cruise portions of the missions when the operator is only monitoring the UAVs as they fly to the battle zone. At the other extreme are situations when the operators are

selecting targets and tending to vehicle malfunctions in time critical portions of the mission. In order to characterize the effects of fatigue on the range of task performance and psychophysiological consequences, the current study employed three tasks. The cognitive demands of these tasks ranged from a simple reaction time task to a four part simulated aviation task to a simulated UAV mission.

METHODS

Subjects

Nine young adults (8 males and 1 female) served as subjects after giving informed consent. Their mean age was 25 years, range 22 to 36 years. Prior to the study, all of the participants were reportedly on a normal daytime schedule in which they generally reported to work between 0700 and 0800 and worked until 1600 or 1700. According to actigraph data, the participants acquired an average minimum of 7 hours and 52 minutes of sleep on the night prior to the beginning of any of the sleep deprivation periods.

Materials and Procedures

A variety of assessments were conducted in an effort to characterize the global impact of fatigue on performance. Subjects were trained on all tasks prior to the sleep-deprivation period to minimize practice effects.

Performance Tasks

Psychomotor Vigilance Task (PVT). The level of vigilant attention was assessed with the PVT (Dinges et al., 1997). This task required subjects to hold a small device equipped with an LED digital display, and to respond to the onset of a digital counter by pressing either of two response buttons as soon as the stimulus appeared. The response, which stopped the stimulus counter,

displayed reaction time (RT) in milliseconds for a 1-second period. The inter-stimulus interval varied randomly from 2 s to 10 s, and the task duration was 10 minutes (which yielded approximately 80 RTs per trial). Data from this test included: mean RT, standard deviation (SD) of the RTs, median RTs, SD of the median RTs, the reciprocal RTs, the number of reaction times greater than 500 milliseconds (lapses), the square root transformation of the lapses, the mean of the slowest 10% of the RTs, the SD of the slowest 10% of the RTs, and the overall reaction time.

<u>Multi-Attribute Test Battery (MATB).</u> The MATB (Comstock and Arnegard, 1992) is a computerized aviation simulation test that required participants to perform an unstable tracking task while concurrently monitoring warning lights and dials, responding to computer-generated auditory requests to adjust radio frequencies, and managing simulated fuel flow rates using various key presses. This test was controlled by a personal computer equipped with a standard keyboard, joystick, and mouse. Data on tracking errors, response times, time-outs, false alarms, and accuracy rates were calculated.

<u>Operator Vehicle Interface Task (OVI).</u> This task required the subjects to visually monitor the progress of four autonomous vehicles as they flew a preplanned bombing mission. The mission consisted of three intermixed components, a cruise portion during which the vehicles flew from waypoint to waypoint, an easy target condition, and a difficult target evaluation condition. Four threat areas were assigned to each vehicle for a total of sixteen radar images (SARs) to be evaluated per mission. When the vehicles reached designated waypoints, the radar images were

automatically captured, subjects were then required to give commands to download and view the SAR image of the target area. The subjects then had to find and designate six targets by a pre-set time before the vehicle reached the weapons release waypoint. Three categories of targets were used and the subjects were required to use a predetermined set of priorities when selecting targets (see Figure 1, left). Because the entire SAR image could not be viewed at one time (Figure 1, right), the subjects had to pan around the image to locate the targets. Following target designation, a weapons release command had to be given before the vehicle reached the weapons release waypoint. If the release command was not given before the vehicle reached the release waypoint, the bombs from that vehicle could not be released thereby greatly reducing the effectiveness of the entire mission for that vehicle. SAR images were presented at two levels of complexity. The more difficult contained a larger number of distracters and required more complex decisions concerning target priority. The occurrence of the eight easy and eight difficult SAR images was mixed and each mission required 25 minutes to complete. Simultaneously, the subjects monitored the status of each vehicle by observing messages showing potential vehicle problems such as fuel pump failures. The subjects' memory load was manipulated by having them keep two aircraft problem combinations in memory until a command was given which signified which malfunction had reached a critical level and had to be corrected. The subjects then selected the appropriate vehicle from a pull down menu and using other pull down menus found and selected the appropriate fix for the indicated vehicle problem.

The number of designated mean points of impact (DMPI) that were placed (i.e., the number of targets designated), the number of targets hit, the number of non-targets designated (false alarms) and whether or not the command to release the weapons was executed in time were recorded. The vehicle status task was scored by the number of correct solutions, the number of time-outs and the reaction time for responding to a critical malfunction.

Subjective Scales

Profile of Mood States (POMS). Subjective evaluations of mood were made with the POMS (McNair, Lorr, and Droppleman, 1981). The POMS is a 65-item questionnaire which measures affect or mood on 6 scales: 1) tension-anxiety, 2) depression-dejection, 3) anger-hostility, 4) vigor-activity, 5) fatigue-inertia, and 6) confusion-bewilderment. Scores on each scale were analyzed to determine fatigue effects.

Visual Analog Scales (VAS). In addition to the POMS, subjective sleepiness and alertness were measured via the VAS (an adaptation of the version developed by Penetar et al., 1993). This questionnaire consists of several 100-millimeter lines, each labeled at the left end with the words "not at all" and the right end with the word "extremely." Centered under each line were the test adjectives as follows: "alert/able to concentrate," "anxious," "energetic," "feel confident," "irritable," "jittery/nervous," "sleepy," and "talkative." The participants indicated the point on the line that corresponded to how he/she felt along the specified continuum at the time at which the test was taken. The score for each item consisted of the number of millimeters from the left side of the line to the location at which the participant placed the mark.

Psychophysiological measures

Electroencephalographic (EEG) data. EEG data were recorded with gold plated cup electrodes that were attached to the scalp and both mastoids with collodion at the following 10/20 electrode sites: F7, Fz, Cz, Pz, and Oz. One mastoid served as reference and the other as ground. Eye and cardiac activity were recorded using disposable Ag/AgCl electrodes. The electrooculography (EOG) electrodes were placed above and below the right eye for vertical movement and blink activity. Electrodes placed next to the outer canthus of each eye were used to record horizontal ocular activity. The electrocardiogram (ECG) electrodes were placed on the sternum and on the left clavicle. All of the psychophysiological data were amplified and digitized at 200 Hz with Cleveland Biomedical BioRadio 110 telemetry units. The bandpass was from 0.5 Hz to 52.4 Hz. The digitized data were stored on a computer disk and simultaneously reduced on-line with a laboratory developed software program, NuWAM (Krizo, Wilson & Russell, 2005). Eye artifacts in the EEG data were corrected using an adaptive filter with inputs from the vertical and horizontal eye channels (He, Wilson & Russell, 2004). The corrected EEG and the EOG data were submitted to a fast Fourier transformation (FFT) every second. Interbeat intervals were calculated, on-line, from the ECG data. The EEG data were separated into five bands for further statistical analysis. The bands were: delta - 2.0 to 4.0 Hz, theta - 5.0 to 8.0 Hz, alpha - 9.0 to 13.0 Hz, beta - 14.0 to 32.0 Hz and gamma - 33.0 to 43.0 Hz.

Outliers in the EEG data were identified using the JMP software statistical package (SAS Institute Inc, Cary, NC, USA).

The mean and SD of the reduced data for each condition and variable were calculated and those data which were 2 standard deviations from the mean were identified. Experience in our laboratory has shown this to be a conservative method to identify artifacts in the data. These outliers were excluded from subsequent statistical analysis.

Cardiac measures. The R waves from the ECG data were located and the interbeat intervals were calculated. The interbeat intervals were examined and corrected for extra and missed beats. The corrected data were used to determine mean interbeat intervals and heart rate variability using the PBfilter (Delta-Biometrics, Inc. Bethesda, MD). Two bands were used, Taube Herring Mayer (THM) band from 0.06 to 0.14 Hz and the respiratory sinus arrhythmia band (RSA) from 0.15 to 0.25 Hz.

Oculography data. An EyeLink II System (SR Research Ltd., Ontario Canada) video based head mounted eye tracking was used to measure the pupil area. Two eye cameras, with built-in illuminators, allowed for binocular pupil area measurement at 250 Hz.

Wrist activity monitors.

Wrist monitors (Ambulatory Monitoring, Inc., Ardsley, NY) were used to determine the amount of sleep obtained during the night prior to reporting to the laboratory. Computer-generated actigraphs were analyzed to verify that participants had obtained a minimum of 7 hours of sleep the night prior to reporting for testing. These actigraphs also were used to ensure that subjects did not nap from the time at which they awakened (in the morning prior to the night of sleep deprivation) until the time at which they reported to the laboratory.

Testing schedule

Prior to the actual sleep-deprivation, all subjects were trained on the tasks in order to reduce potential confounds attributable to practice effects. Several days prior to testing the subjects practiced the OVI task three times for approximately one hour each session. On the day immediately prior to sleep deprivation, training on all tasks began at 1200 and ended at approximately 1730. Participants completed six iterations of the MATB, five of the OVI, two POMS, two VAS, and one PVT. At the conclusion of the training session participants donned a wrist activity monitor and were asked to wake up at 0600 the next day (or 0700 if necessary to obtain the requisite hours of sleep). They returned to the testing facility at 1900. Upon reporting, the electrodes were attached. Each EEG and mastoid placement site was cleaned with acetone and the electrodes were attached with collodion and then filled with electrolyte gel. Disposable, pre-gelled, self-adhesive electrodes were used for the ECG and EOG sites. Prior to testing, impedances were reduced to less than 5000 Ohms at each EEG and mastoid electrode and to less than 10,000 Ohms at each EOG and ECG electrode.

The participants then proceeded to the first test session which was a pre-deprivation session that began at 2100 with the MATB. During the MATB, EEG, EOG and ECG data were recorded continuously. Then, at 2205 the participants completed a resting eyes-open/eyes-closed EEG while seated at the OVI test station (4 minutes total). This was followed by the OVI task beginning at 2210. For each of the OVI runs EEG, EOG and ECG activity were recorded continuously. Also, during the second and fourth OVI tests (at 0110 and 0410), the eye tracking device was used to

record pupil area. Following the OVI task at 2240, the participant completed the PVT, in which EEG, EOG, and ECG activity were recorded continuously. Finally, at 2255 the POMS and VAS were completed to conclude the test session. Afterwards the participant had an hour break before beginning the next test session.

During the rest of the sleep-deprivation cycle, each task was begun three hours after the beginning of the previous run. Overall, the participants completed five test sessions (starting at 2100, 0000, 0300, 0600, and 0900) and the last of these sessions ended at 1115, after 28-29 hours without sleep (the actual length of the wakefulness period was dependent on the exact wakeup time that was necessary to ensure the volunteer acquired 8 hours of pre-study sleep).

While in the testing facility, meals and snacks were provided as were video games and movies. Each participant was continuously monitored from the time of reporting until departing to ensure that involuntary sleep episodes did not occur.

At the conclusion of the deprivation period, the participant's electrodes were removed; he/she was debriefed and then driven home by a staff member or a family member. Participants were cautioned that they should not drive, operate complex machinery, or engage in other potentially dangerous tasks until obtaining at least one full night of normal sleep.

Data Analysis

Analysis of variance (ANOVA) was used to statistically evaluate the performance, psychophysiological and subjective data. Paired comparisons, t-tests, were performed to determine significant differences following significant ANOVAs using p =< 0.05.

RESULTS

Performance Data

PVT

The number of correct responses demonstrated a significant effect associated with the time of testing ($F(4, 32) = 3.33$, $p = 0.022$), see figure 2 (left). At 0740, the number of correct responses was significantly lower than at the other four data collection times. The median of the correct reaction times and mean of the reciprocal reaction times (RRT) were affected by prolonged wakefulness, ($F(4, 32) = 9.33$, $p < 0.0001$ and $F(4, 32) = 10.65$, $p < 0.0001$, respectively). The median RT was significantly longer at 0740 than at the other four testing times, and the median RT at 2240 was significantly shorter than at the other four sessions, see figure 2 (right). The mean RRT was shortest at 0740 and longest at 2240 in comparison to the other four times. The lapses greater than 500 ms and the square root transformation of the lapses were also significantly affected by time of testing ($F(4, 32) = 4.97$, $p = 0.0031$ and $F(4, 32) = 6.19$, $p = 0.00008$, respectively), see figure 2 (bottom). Post hoc tests showed that there were significantly more lapses (and square-root-transformed lapses) at 0740 than at all other testing times. Similarly, the mean and standard deviation of the slowest ten percent of the reaction times were significantly larger at 0740 than at the other

times ($F(4, 32) = 7.28$, $p = 0.00003$ and $F(4, 32) = 4.22$, $p = 0.0075$, respectively).

MATB

The reaction times to simulated warning lights were significantly affected by time of day ($F(4, 32) = 7.43$, $p = 0.0002$), figure 3 (left). The post-hoc tests demonstrated that the RTs at 0600 and 0900 were significantly longer than those collected at the other test times, while not significantly different from each other. All of the other comparisons were significantly different. The standard deviation of these reaction times also were significantly affected by time of testing ($F(4, 32) = 6.29$, $p = 0.00008$). Deviations were larger at 0600 than at 2100, 0000, 0300 and 0900 hours while the standard deviations at 2100 were smaller than all other data collection times except 0000. There was a time-of-testing main effect on RMS tracking errors as well ($F(4, 32) = 12.46$, $p < 0.0001$) with errors at 0600 and 0900, while not different from each other, being greater than those found at 2100, 0000 and 0300, figure 3 (right). The other paired comparisons were also significantly different except for the 2100 and 0000 comparison.

OVI

The number of weapons release waypoints successfully completed was significantly higher for low than high difficulty SARS ($F(1, 32) = 49.98$, $p<0.0001$) and there was an interaction between workload and hours awake ($F(4, 32) = 6.21$, $p = 0.0006$), see figure 4 (left). Separate ANOVAs for the low and high difficulty SAR conditions showed that performance of the low difficulty portions was significantly affected by testing time ($F(4, 32)= 13.35$, $p<0.0001$) whereas the performance of the high difficulty portions

was not. In the low-workload condition, fewer weapon release waypoints were successfully made at 0110, 0710 and 1010 than at 2110 and 0410 hours. Conversely, although the number of false alarms (Figure 4, right) was significantly affected by the time of testing ($F(4, 32) = 3.56$, $p = 0.0165$), workload ($F(1, 32) = 68.16$, $p < 0.0001$) and the interaction of testing time and workload ($F(4, 32) = 3.87$, $p = 0.011$), there were no false alarms in the low difficulty condition versus an increase in false alarms at 2110 and 1010 in the high-workload condition. The number of DMPIs placed significantly varied only as a function of workload ($F(1, 32) = 6.67$, $p = 0.014$) with more DMPIs being placed during the low difficulty SAR condition than during the high difficulty condition. For the VHT task, correct response reaction times were affected by testing time ($F(4, 32) = 3.31$, $p = 0.02$) with the shortest RTs at 2110 and the longest at 1010 (see figure 4, bottom). The RTs at 2110 were significantly shorter than those collected at the other four testing sessions and the RTs at 1010 were longer than those recorded at 2110, 0410 and 0710.

Subjective Data

POMS

The POMS fatigue scale was significantly affected by time of testing ($F(4, 32) = 22.74$, $p < 0.0001$), with the largest fatigue ratings at 0755 compared to those at the other four testing times. All of the other comparisons were significantly different except for the 0455 versus 1055 comparison (see figure 5, top left). The vigor scale was also significantly affected by time of testing ($F(4, 32) = 20.88$, $p < 0.0001$). The rating at 0755 was lower than at any of the other times, and all other comparisons were significantly different except for the 0455 versus 1055 test

(figure 5, top right). The confusion scale of the POMS was affected by time of day ($F(4, 32) = 14.48$, $p < 0.0001$), and in this case, scores at 0755 were higher than at all other testing times except for 1055. As with the other two previous scales, all other comparisons were significantly different except for the 0455 and 1055 testing times (figure 5, bottom). The tension/anxiety, depression/dejection, and anger scales were not significantly affected by the time of testing.

VAS

The subjective reports of alertness were significantly affected by the time of testing ($F(4,32) = 33.19$, $p < 0.0001$), with the 0755 and 1055 scores, while not different from each other, being lower than the scores obtained at all other testing times (figure 6, top left). All other comparisons were also significantly different. Energy scores ($F(4, 32) = 15.77$, $p < 0.0001$), confidence ratings ($F(4, 32) = 5.42$, $p = 0.0019$), and talkativeness ($F(4, 32) = 16.02$, $p < 0.0001$) also were impacted by the number of hours awake (i.e., testing time). See figure 6, top right, bottom left, and bottom right, respectively. Energy ratings at 0755 were lowest, and all other session comparisons were significantly different except for those between 1055 and 0155 and between 1055 and 0455. Confidence ratings at 0755 were significantly lower than those at all other testing times, with all other comparisons again significantly different except for 1055 which was not statistically different from 0155 and 0455. In addition, 0155 was not different from 0455. Talkativeness at 0755 was lowest, and all other comparisons again were significantly different except for those between 1055 and 0155 and between 1055 and 0455. The sleepiness scale was a virtual mirror image of the

previously-mentioned scales in that the significant time effect ($F(4,32) = 21.36$, $p < 0.0001$) was due to the highest ratings at 0755 in comparison to the other sessions. All other comparisons were significantly different except the ones between 0155, 0455 and 1055. The anxiety, irritability and jittery scales were unaffected by sleep loss.

Psychophysiological Data

Electroencephalographic Data

Resting Condition. Examination of the EEG log power ANOVA results showed that the time of day effect was statistically significant for all electrode sites in the delta, theta and alpha bands, see figure 7. EEG power increased in the delta and theta bands over the five testing sessions with significant increases at the 0705 and 1005 testing sessions, see figure 8. The alpha band power decreased over time, see figure 9. There were significant interactions between time of testing and eyes open or closed at all five electrode sites in the alpha band. The decrease in alpha band power was primarily seen in the eyes closed condition while the power during the eyes open condition was fairly constant after an initial drop following the first testing period. There were significant differences between eyes open and eyes closed conditions for the delta band at electrodes T5, Cz and Pz and for the alpha band at Pz and Oz. The eyes closed condition exhibited larger EEG power for both bands than the eyes open condition.

PVT. The ANOVA of the EEG log power data collected during the PVT task performance showed significant increases in the delta band at electrodes F7, Pz and Oz ($F(4, 24) = 4.12$, $p = 0.011$; $F(4, 23) = 5.64$, $p = 0.003$; $F(4, 24) = 6.39$, $p = 0.001$, respectively).

The ANOVA analysis of the theta band EEG power at F7, Cz and Pz sites showed that there were significant effects due to the time of testing ($F(4, 24) = 4.01$, $p = 0.012$; $F(4, 24) = 3.02$, $p = 0.038$; $F(4, 23) = 3.13$, $p = 0.034$, respectively). The peak EEG power was found at the 0740 testing session, see figure 10. The delta power at the 0740 testing session was significantly larger than the power at 2240, 0140 and 0440 at F7, Pz and Oz. Also, the power was significantly larger at 0740 than at 1040 at the Pz and Oz sites. The EEG delta band power was significantly larger during the 1040 than at the 2240 testing session at F7 and Oz. The theta band results showed that the power during the 0740 session was significantly greater than at the 2240 and 0140 at Cz and Pz. The theta band power at 0740 was also significantly larger than at 1040 at the Cz and Pz sites. At F7, the theta power at 0440 was significantly larger than that found at 2240.

MATB. The EEG collected during MATB task performance showed significant differences due to time of testing in the log power at Pz in the alpha band ($F(4, 31) = 3.00$, $p = 0.033$), see figure 11. The 0600 and 0900 testing sessions were associated with greater alpha band power than the power found at 0000; the alpha band power at 0900 was significantly larger than at 0300. In the beta band there were significant differences in power at sites F7, Cz and Pz, see figure 11. The beta band power at the 0600 testing session was significantly larger than the power at 2100 and 0000 at all three electrode sites. Further, beta power at 0600 was significantly larger than at 0300. The beta band power during the 0900 session was significantly larger than that at the 2100 session at the F7 and Cz sites. There were significant changes in the gamma band power due to hours awake during the MATB task

performance at F7, Cz, Pz and Oz ($F(4, 32) = 3.80$, $p = 0.012$; $F(4, 32) = 3.22$, $p = 0.025$; $F(4, 31) = 3.81$, $p = 0.012$, $F(4, 32) = 3.49$, $p = 0.018$, respectively). Significantly larger gamma band EEG power was found at all four sites at the 0600 session when compared to the 2100 and 0000 sessions. At Pz at 0600 power was also greater than at the 0300 at the Pz site. Further, the gamma band power recorded at 0900 was significantly larger than the power at 2100 at all four sites.

<u>OVI</u>. During OVI performance, the time of data collection produced significant changes in only the delta band power at Pz ($F(4, 31) = 2.98$, $p = 0.034$) and Oz ($F(4, 32) = 3.30$, $p = 0.023$) (figure 12) with a significant time-of-testing by task difficulty interaction at Oz ($F(8, 64) = 2.12$, $p = 0.046$). At Pz the delta band power at 0410 was significantly smaller than at the 2210 and 1010 with the power at 1010 significantly larger than at 0710. At the Oz electrode site only the low condition was significantly effected by time of testing ($F(4, 32) = 5.12$, $p = 0.003$). At the 0410 testing session the delta power was significantly lower than at the other four testing sessions.

The effects of the task difficulty produced significant differences in the delta band at F7, T5, Pz and Oz, ($F(2, 16) = 19.63$, $p = 0.001$; $F(2, 16) = 9.56$, $p = 0.002$; $F(2, 16) = 14.32$, $p = 0.001$; $F(2, 16) = 17.40$, $p = 0.001$, respectively) figure 12. At all five electrode sites the cruise condition was associated with greater delta band power than both the low and difficult SAR conditions. The low and difficult conditions were not significantly different. Further, significant differences due to

task difficulty were found in the theta band at Cz ($F(2, 16) = 4.07$, $p = 0.037$). The high difficulty condition showed greater theta band power than the cruise condition. There were also significant differences in the beta and gamma bands at F7, Cz and Pz, ($F(2, 16) = 4.85$, $p = 0.023$; $F(2, 16) = 10.52$, $p = 0.01$; $F(2, 16) = 10.27$, $p = 0.001$, respectively), figure 13. These differences were the result of reduced power during the low and high SAR conditions compared to the cruise condition.

Heart Rate Data

Resting Condition. Neither the heart rate, THM nor RSA data showed any significant differences due to time of testing.

PVT. None of the comparisons for the PVT were significantly different.

MATB. Both measures of heart rate variability were significantly affected by the time of testing during the MATB task performance (THM, $F(4, 32) = 14.07$, $p < 0.0001$; RSA, $F(4, 32) = 7.04$, $p = 0.003$). The variability increased as the testing progressed, see figure 14. The variability at 0600 and 0900 was significantly greater than at the other three testing times while 0600 and 0900 were not significantly different. The results showed that the variability at 0300 was significantly higher than at 2100 and 0000.The interbeat intervals were not significantly affected by time of testing.

OVI. The interbeat intervals showed significant effects due to time of testing ($F(4, 32) = 5.53$, $p = 0.002$). The interbeat intervals found at the 0410 testing session were significantly larger than those at the other four testing times. The two measures of heart rate variability recorded at the five testing sessions were not significantly different.

Pupil area

Pupil area data recorded at the 0110 and 0710 OVI testing sessions were not significantly different for either the right or left pupil measures. However, both the right and left pupil areas were significantly affected by OVI task difficulty ($F(2, 16) = 11.38$, $p = 0.0008$; $F(2, 16) = 5.21$, $p = 0.018$, respectively), These comparisons included the cruise, low and high difficulty conditions. Post hoc comparisons revealed that pupil area significantly increased from cruise to low difficulty and also significantly increased from the low to high difficulty conditions while the subjects performed the OVI task.

DISCUSSION

One night's sleep deprivation affected some but not all aspects of task performance on the PVT, MATB, and OVI. In many cases, the psychophysiological data (primarily EEG) collected during the performance of each task and during a resting condition generally paralleled the performance changes, as did many of the subjective indicators of well being. The timing of the significant task degradation effects was somewhat unique for each of the three tasks.

The simple reaction time task, PVT, exhibited the longest reaction times with the most variability and the most response lapses at the next-to-the-last test session (at 0740), after approximately 25 hours of continuous wakefulness. This is consistent with earlier reports of increased performance irregularities as a function of sleep loss and circadian influences (Dinges, 1990; Moore-Ede, 1993). Although there were improvements in PVT performance towards the end of the study (at 1040), this was likely due to an "end-spurt" effect rather than

any sort of physiological recovery since the subjects were aware this was the final testing session. The EEG collected while subjects were performing the PVT showed increased lower frequency delta and theta activity at the 0740 testing session. This was to be expected since increases in slow-wave EEG have previously been associated with decreased alertness (Wright and McGown, 2001). Subjective measures of well being were similarly affected in that ratings of fatigue, confusion, and sleepiness showed the greatest increases at 0755, while measures of vigor, alertness, energy, and talkativeness showed the greatest decreases at this time (self-ratings of confidence were lowest at 0455). Once most of these mood ratings deteriorated, they tended to remain relatively degraded for the remainder of the study.

Performance on two of the four tasks in the more complex MATB, showed similar decrements between the third and fourth (next-to-last) testing sessions as were observed in the PVT. However, reaction times to MATB warning lights and MATB tracking errors, revealed no end-spurt improvement during the final test administration. Once performance declined, it remained impaired until the end of the sleep-deprivation period. Although the MATB is a more difficult task overall, it is noteworthy that the only two tasks which showed statistically-significant decrements were the ones that required fairly simple responses (reacting to warning lights) or continuous monitoring and motor output (vigilantly completing the tracking task). The other two non-degraded tasks were the communications task which required more complex input and output processing and the resource management task which required the development and execution of a strategy. Such differences may be due to the fact that very simple tasks

tend to be less interesting and less engaging than more complex tasks, which can make such tasks more vulnerable to the effects of sleep loss (Wilkinson, 1964). With regard to physiological correlates of task performance, the EEG and heart-rate measures collected during the MATB were highly correlated with the performance effects. There was increased power in the higher frequency beta and gamma EEG bands and concurrent increases in heart-rate variability during the last two test sessions. The expected elevations in slow-wave EEG, observed under resting conditions and during the performance of the PVT, did not occur. This may be because performing the more engaging and complex MATB task (considering the requirement to perform 4 subtasks simultaneously) overcame the fatigue effects of increased lower frequency enhancement as seen in the PVT task and produced the increased higher frequency EEG activity. Further, the finding of impaired self-reported mood states observed near the MATB testing times supports the contention that fatigue from progressive sleep loss was hampering the subjects' abilities to perform this task. Self-reported mood status was the worst at 0755 (as noted above in the description of the PVT results), but self-rated mood also was degraded at 0455 and 1055—the times which bracketed the impaired MATB sessions.

The results for the most complex task, the OVI, are not as straight forward as those observed for the PVT and MATB. Although the number of DMPIs placed varied as a function of workload, there were no effects attributable to sleep loss. However, three other measures were affected by the combination of both workload and fatigue, albeit in different ways. Only in the low-workload condition was the number of completed weapons release points

significantly affected by fatigue, whereas only during the high difficulty portion of the task was the false-alarm rate significantly altered. However, successful weapon releases during the low-workload condition declined from 2110 to 0110, improved from 0110 to 0410, and then declined once again at 0710 and 1010. Thus, under the low workload condition, this aspect of OVI performance was quite variable despite a linear increase in sleep pressure. Conversely, the false-alarm rate was highest during the first session (at 2110) after which it declined during the middle sessions (0110, 0410, and 0710) before once again increasing at the end of the sleep-loss cycle. Perhaps the greater number of false alarms at the outset might have been due to a learning or warm-up phenomena while those at the last session were due primarily to an increase in fatigue (having been awake for approximately 28 hours); however, the notion that practice effects accounted for the poorer performance prior to sleep loss (at 2110) is complicated by the fact that a similar overall pattern was not observed in the weapons-release data where performance at the outset was better than performance at the end. Nonetheless, it should be noted that in both cases, performance was significantly degraded at the end of the sleep-deprivation period in comparison to performance at one or more points earlier in the testing cycle, and this makes it quite likely that increased fatigue was responsible. This is consistent with the effect observed on the Vehicle Health Task in which the longest reaction times clearly occurred at the end of testing (at 1010) whereas performance was much better at the outset (at 2110 with response accuracy being maintained throughout) (Falleta et al., 2003). The idea that fatigue-related difficulties were responsible for these last-

session decrements on three OVI performance variables is further bolstered by examining both the resting EEG data (which preceded each OVI) and the EEG data collected during each iteration of the OVI. In both cases, delta activity was greater at approximately 1000 than during one or more of the previous testing times. Nevertheless, the absence of consistent task effects across all of the performance measures makes it impossible to directly explain all of the OVI findings with a straightforward fatigue (or circadian) interpretation. No doubt, the interaction between the effects of fatigue and the impact of cognitive load is complex, but this finding is, in and of itself, important. In fact, it rather clearly shows that the *type of task to be performed* could be as important as *the degree of sleep loss prior to task performance* in predicting the ultimate probability of operator success—a notion which is consistent with earlier work published by Wilkinson (1964). The task difficulty effects (workload) persisted across the five testing sessions with the differences primarily between the cruise and the combined low and high difficulty conditions. This was correlated with the widespread distribution of effects over the scalp in the delta, beta and gamma bands and the more localized theta effects at the Cz electrode.

In terms of the central and peripheral physiological data collected in this study, the typical power increases in the lower frequency bands of delta and theta, with the accompanying decrease in alpha band power, revealed that sleep loss was progressively compromising operator status. These effects increased when the testing conditions were more soporific (under conditions of eyes closed versus eyes open). Thus, in general terms, the central

nervous system (EEG) data supported the performance and subjective mood-state findings. The peripheral measures (HR and HRV) were somewhat less clear-cut in that changes were observed only during the MATB and OVI task performance, while there were no significant time of day effects during the resting test session or during PVT task performance. Interestingly, during the MATB, HRV increased as a function of sleepiness while it might have been expected to decrease if the task performance required greater cognitive resources because of the fatigue effects (Mulder, Mulder, Meijman, Veldman & Roon, 2000). It appears that the increase in HRV associated with sleep loss is the stronger effect. Further, the increased HRV in both bands paralleled the changes in the MATB by exhibiting the largest effects at the last two testing sessions. The heart rate slowed significantly only during the third OVI testing session.

The pupil area measure was not affected by the sleep loss as might have been expected based on earlier findings published by Stern and Ranney (1999) and Caldwell et al. (2003). However, there were significant increases in pupil area with increased task difficulty in both testing sessions where pupil area was recorded. This is consistent with a large body of literature demonstrating increased pupil diameter with increased cognitive task loads (for a review, Sirevaag & Stern, 2000). One difficulty with interpreting the pupil results is the possibility that the light levels from the OVI screen during the cruise, low and high difficulty conditions may have been sufficiently disparate to cause the differences in pupil diameter. However, the results are consistent with studies which have held luminance levels constant while manipulating the cognitive difficulty of tasks (Sirevaag &

Stern, 2000). Even though fatigue has been associated with pupil diameter decreases it is possible that the opposing pupillary dilation effects of cognitive task difficulty overcame the fatigue effects and produced the resulting lack of significant changes due to time of testing.

In summary, the data from the present investigation show that fatigue from moderate extended wakefulness generally impairs operator readiness, but that the degree of impairment is at least partially dependent on the task to be performed and the outcome variable under scrutiny. Tasks which are more engaging are likely to show fewer (or at least less-predictable) decrements than simpler, less engaging tasks (Wilkinson, 1964). All of the subjects spontaneously reported during the debriefing session that the OVI task was more engaging than the PVT and MATB tasks, and interestingly, decrements on the OVI task were not as consistent as those on the MATB, and clearly not as uniform and robust as those seen on the PVT. It is possible that within the OVI task the greater impact of sleep loss on the easy OVI task portion was also related to the engagement factor. The performance during the high difficulty portion of the OVI was less affected by fatigue, although the increased tendency towards targeting false alarms toward the end of the sleep-deprivation period may have implications for so-called "friendly-fire" episodes. In general, it appears that simple RT tasks such as the PVT may not be good predictors of performance on more complex, engaging, tasks after one night's sleep loss.

Generalized changes in the EEG data recorded across all of the performance tasks as well as under non-task resting conditions, suggest that it may be possible to design monitoring

systems that would be able to detect fatigue. The psychophysiological data are continuously present and can be unobtrusively collected which makes them candidates for such an application. Psychophysiological signals have been used to detect loss of engagement in the task (Scerbo, 1996) and changing levels of cognitive workload (Wilson, & Russell, 2003). Further, this information has been used to close the loop in an adaptive aiding system to improve performance (Freeman, Mikulka, Prinzel, & Scerbo, 1999; Wilson, & Russell, 2004). Such systems could be used to detect fatigue and to alert the operator and/or to notify the system or other personnel so that corrective actions could be taken.

APPLICATIONS

Fatigue effects can produce impaired performance. The degree of performance impairment seems to be a function of the numbers of hours awake and the "engagement" value of the task. Simple, nonengaging tasks, such as the PVT, tend to show performance decrements sooner than more complex and more engaging tasks such as the MATB and the OVI task. When making determinations with regard to fatigue it is well to consider the nature of the task that operators will be performing.

References

Akerstedt, T. (1995). Work hours, sleepiness and the underlying mechanisms. *Journal of Sleep Research*, 4(suppl. 2), 15-22.

Akerstedt, T. (2004). Mental fatigue, work and sleep. *Journal of Psychosomatic Research*, 57, 427-433.

Bourgeois-Bougrine, S, Carbon, P, Gounelle, C, Mollard, R, & Coblent, A. (2003). Perceived fatigue for short- and long-haul flights: A survey of 739 airline pilots. *Aviation, Space and Environmental Medicine*. 74, 1072-1077.

Byrne, E. A. & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology,* 42, 249-268.

Caldwell, J. A., & Caldwell, J. L. (2003). *Fatigue in aviation: A guide to staying awake at the stick*. London: Ashgate Publishing Co.

Caldwell, J. A., Caldwell, J. L., Brown, D. L., Smythe, N. K., Smith, J. K., Mylar, J. T., Mandichak, M. L., & Schroeder, C. (2003). *The effects of 37 hours of continuous wakefulness on the physiological arousal, cognitive performance, self-reported mood, and simulator flight performance of F-117A pilots*. USAF Research Laboratory Technical Report No. 2003-0086. Brooks City-Base, TX: Air Force Research Laboratory.

Caldwell, J. A., Gilreath, S. R., Erickson, B. S., & Smythe, N. K. (2001). *Is fatigue a problem in Army aviation? The results of a survey of aviators and aircrews*. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 2001-03.

Caldwell, J. L. & Gilreath, S. R. (2001). Work and sleep hours of U.S. Army aviation personnel working reverse cycle. *Military medicine*, 166, 159-166.

Co E. L., Gregory, K. B., Johnson, J. M., & Rosekind, M. R. (1999). *Crew factors in flight operations XI: A survey of fatigue factors in regional airline operations*. NASA AMES Research Center Report No. NASA/TM-1999-208799.

Comstock, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research.* Hampton, NY: Langley Research Center, National Aeronautics and Space Administration. NASA Technical Memorandum No. 104174.

Dawson, D., & Reid, K. (1997). Fatigue, alcohol and performance impairment. *Nature*, 388, 235.

Dinges, D. F. (1990). The nature of subtle fatigue effects in long-haul crews. *Proceedings of the Flight Safety Foundation 43rd International Air Safety Seminar; Rome, Italy*. Arlington, VA: Flight Safety Foundation.

Falleti, M. G., Maruff, P., Collie, A., Darby, D. G., & McStephen, M. (2003). Qualitative similarities in cognitive impairment associated with 24 h of sustained wakefulness and a blood alcohol concentration of 0.05%. *Journal of sleep Research*, 12, 265-274.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J. & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, 50, 61-76.

He, P., Wilson, G. F. & Russell, C. (2004). Removal of ocular artifacts from electro-encephalogram by adaptive filtering. *Medical & Biological Engineering & Computing,* 42, 407-412.

Horne, J. A. (1978). A review of the biological effects of total sleep deprivation in man. *Biological Psychology,* 7, 55-102.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress (pp. 279-328). In D. L. Damos, (Ed.), *Multiple task performance*. London: Taylor and Francis.

Krizo, E. J., Wilson, G. F., & Russell, C. A. (2005). NuWAM: The development of a modular, real-time augmented cognition system. *Proceedings of the 11th International Conference on Human-Computer Interaction, Las Vegas, NV.*

Krueger, G. P. (1989). Sustaining military performance in continuous operations: Combatant fatigue, rest and sleep needs (pp 255-277). *Handbook of military psychology*. New York: John Wiley and Sons.

Lamond, N. & Dawson, D. (1999). Quantifying the performance impairment associated with fatigue. *Journal of Sleep Research*, 8, 255-262

Lichacz, F. (2005). Examining the effects of combined stressors on dynamic task Performance, *International Journal of Aviation Psychology*, 15, 45-66.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *Manual for the profile of mood states.* San Diego: Educational and Industrial Testing Service.

Moore-Ede, M. (1993). Aviation safety and pilot error. *Twenty-four hour society* (pp. 81-95). Reading, MA: Addison-Wesley Publishing Co.

Mulder, G., Mulder, L. J. M., Meijan, T. F., Veldman, J. B. P., & van Roon, A. M. (2000) A psychophysiological apoproach to working conditions. In R. W. Backs & W. Boucsein (Eds.), *Engineering Psychophysiology* (pp. 139 - 160). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Perry, I. C. (1974). *Helicopter aircrew fatigue*. AGARD Advisory Report No. 69. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

Rosekind, M.R., Graeber, R.C., Dinges, D.F., Connell, L.J., Rountree, M.S., Spinweber, C.L., et al. (1994). *Crew factors in flight operations IX: Effects of planned cockpit rest on crew performance and alertness in long-haul operations*. Moffett Field, CA: NASA Ames Research Center; 1994. Report No.: DOT/FAA/92/24.

Rosekind, M.R., Co, E.L., Gregory, K.B., & Miller, D.L. (2000). *Crew factors in flight operations XIII: A survey of fatigue factors in corporate/executive aviation operations*. NASA Ames Research Center Report No. NASA/TM-2000-209610.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37-63). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sirevaag, E. J. & Stern, J. A. (2000). Ocular measures of fatigue and cognitive factors. In, R.W. Backs & W. Boucsein (Eds.), *Engineering Psychophysiology* (pp .269-287). Mahweh, N.J.: Lawrence Erlbaum Associates.

Stern, A. & Ranney, T. (1999). *Ocular based measures of driver alertness*. In: Carroll R.J. (ed.) Ocular measures of driver alertness: Technical conference proceedings. FHWA-MC-99-

136. Washington, DC: Office of Motor Carrier and Highway Safety, Federal Highway Administration.

Wilkinson, R. T. (1964). Effects of up to 60 hours' sleep deprivation on different types of work, *Ergonomics*, 17, 175-186.

Wilson, G. F. & Eggemeier, F. T. (1991). Physiological measures of workload in multi-task environments (pp. 329-360). In Damos, D. (Ed.) *Multiple-task performance*. London: Taylor and Francis.

Wilson, G. F. & Russell, C. A. (2003). Operator functional state classification using multiple psychophysiological features in an air traffic control task, *Human Factors*, 45, 381-289.

Wilson, G. F. & Russell, C. A. (2004). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks, *Human Factors*, 45, 635-643.

Wright, N., & McGown, A. (2001). Vigilance on the civil flight deck: Incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters, *Ergonomics*. 44, 82-106.

Xue-liang, C. (2004). An experimental research of mental fatigue using a working memory test. *Chinese Journal of Clinical Psychology*, 12, 331-334.

Author Notes

We thank Lt Jennifer Smith, Cadet Christopher Weimer, Iris Davis, Justin Estepp, Justin Cheuvront and Erin Krizo for help with data collection and analysis and Chuck Goodyear for statistical analysis.

Figure Captions

Figure 1. The three types of targets to be identified, listed in order from lowest to highest priority (left), and an example of the radar image in which target searches were performed (right).

Figure 2. The main effect of testing time (sleep loss) on the number of responses (top left), the median RT (top right), and vigilance lapses (bottom center).

Figure 3. The main effects of testing time (sleep loss) on MATB reaction times to warning lights (left) and RMS tracking errors (right).

Figure 4. The interaction between testing time (hours awake) and task difficulty on the number of SARs successfully completed (top left), the number of targeting false alarms (top right), and the reaction times on the concurrent vehicle health task (bottom).

Figure 5. The effects of sleep loss on POMS fatigue (top left), vigor (top right), and confusion (bottom).

Figure 6. The effects of sleep loss on VAS alertness (top left), energy (top right), confidence (bottom left), and sleepiness (top right).

Figure 7. Significant effects of time for log of power (p • 0.05) by electrode site. If the main effect test of time was significant, a plus sign indicates means increasing over time and a minus sign indicates means decreasing over time. A circle indicates a significant time*condition interaction (only applicable for Resting and OVI).

Figure 8. Theta band power at the Cz electrode site for eyes open and eyes closed during the resting condition.

Figure 9. Pz alpha band power during the resting condition for eyes open and eyes closed.

Figure 10. EEG power from the Pz electrode in the theta band while the subjects performed the PVT task at each of the five testing times.

Figure 11. Alpha band power from the Pz electrode (left) and beta band power at Cz (right) for the five testing sessions while the subjects performed the MATB task.

Figure 12. The effects of sleep loss (testing time) and OVI task difficulty on EEG delta recorded from the Oz electrode site.

Figure 13. The effects of sleep loss on EEG gamma power recorded from electrode Cz while subjects performed the OVI task.

Figure 14. Mean THM band variance during MATB performance for each of the testing sessions.

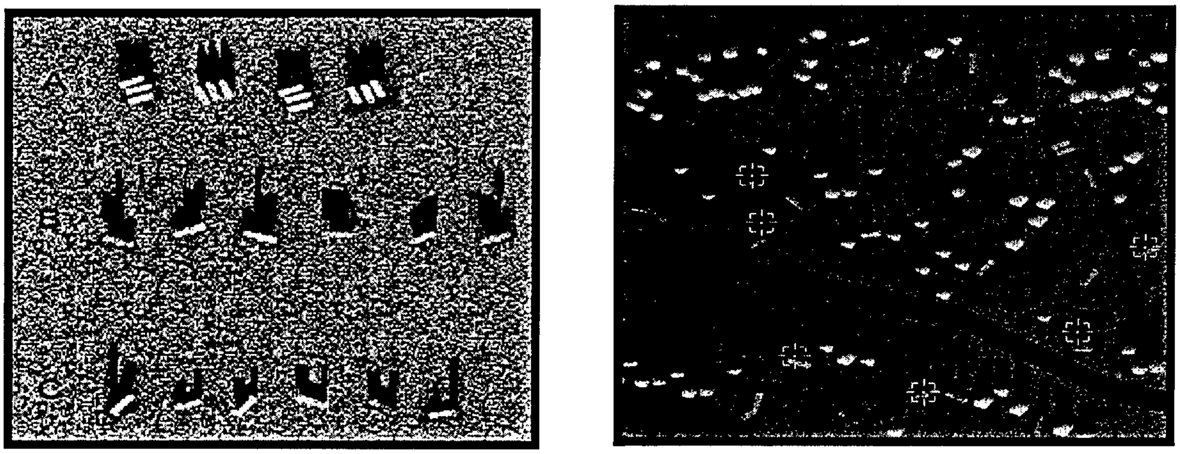


Figure 1.

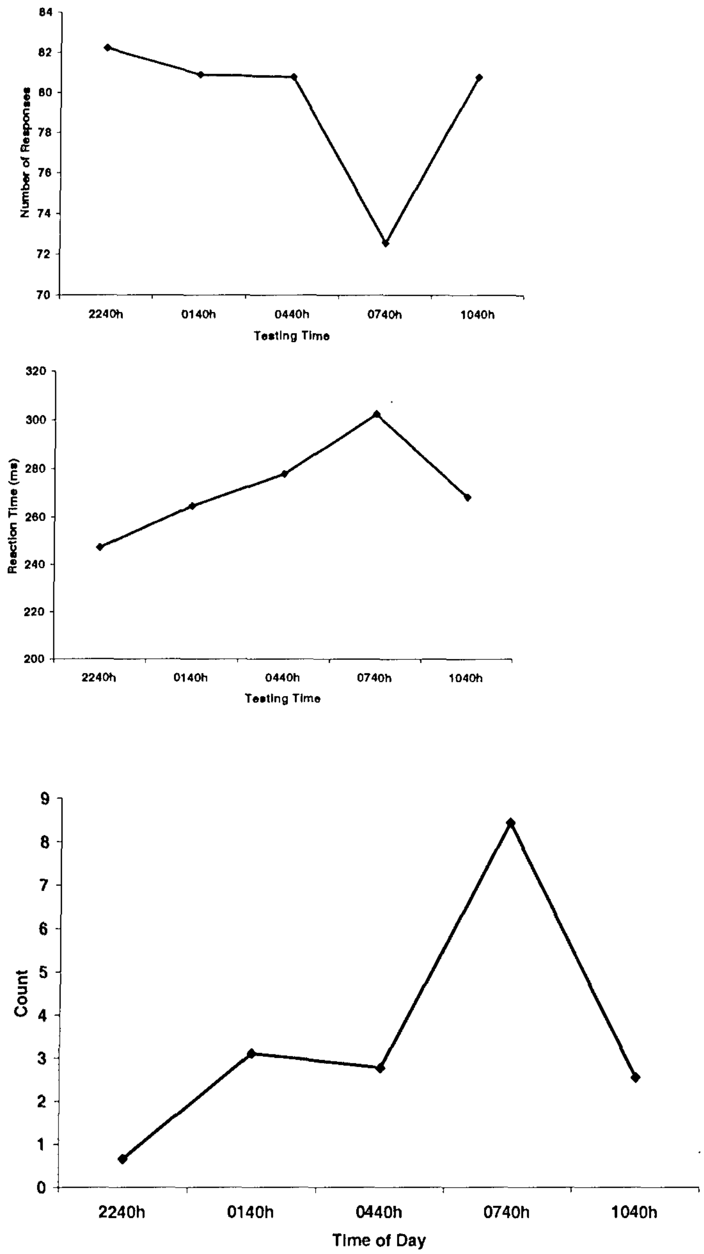

84
82
80
78
76
74
72
70
Number of Responses
2240h
0140h
0440h
0740h
1040h
Testing Time
320
300
280
260
240
220
200
Reaction Time (ms)
2240h
0140h
0440h
0740h
1040h
Testing Time
9
8
7
6
5
4
3
2
1
0
Count
2240h
0140h
0440h
0740h
1040h
Time of Day

Figure 2.

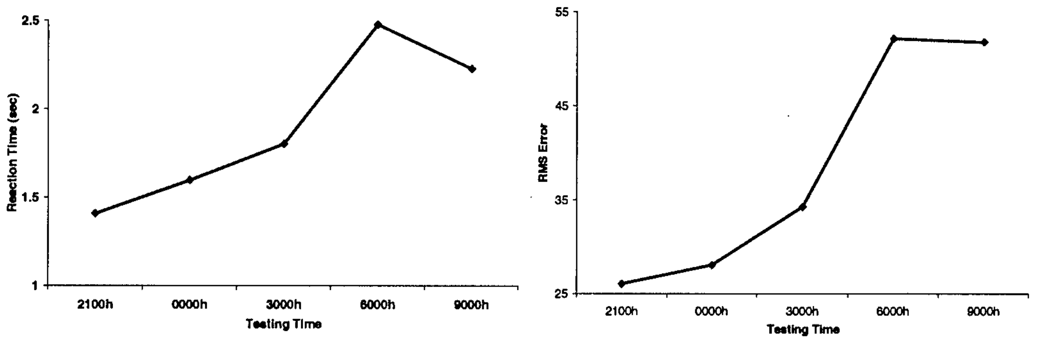

2.5
2
1.5
1
Reaction Time (sec)
2100h
0000h
3000h
6000h
9000h
Testing Time
55
45
35
25
RMS Error
2100h
0000h
3000h
6000h
9000h
Testing Time

Figure 3

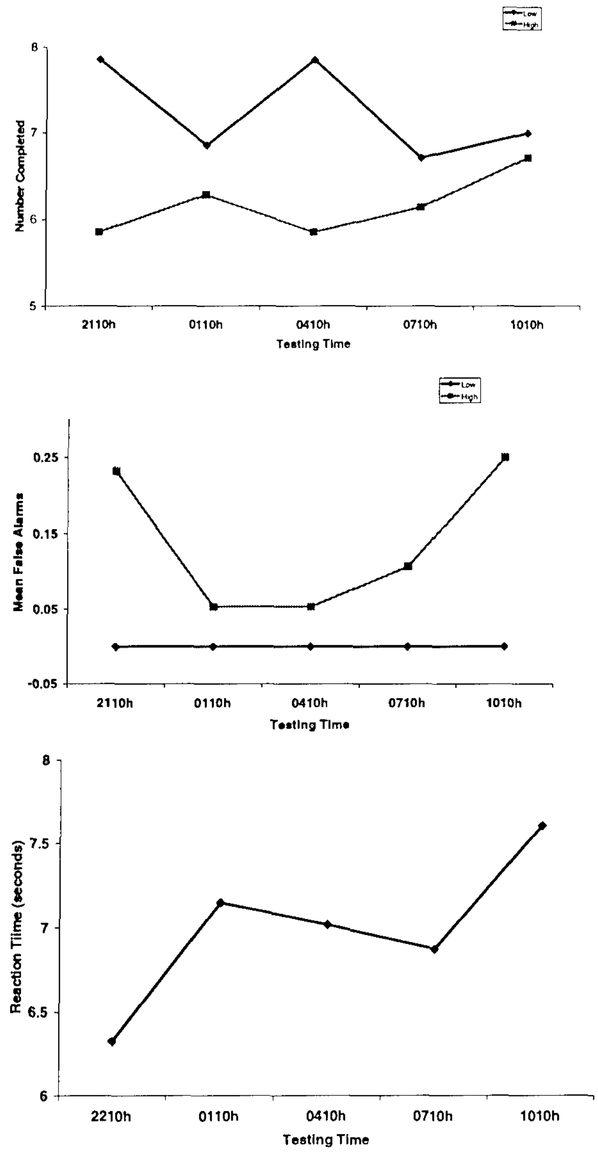
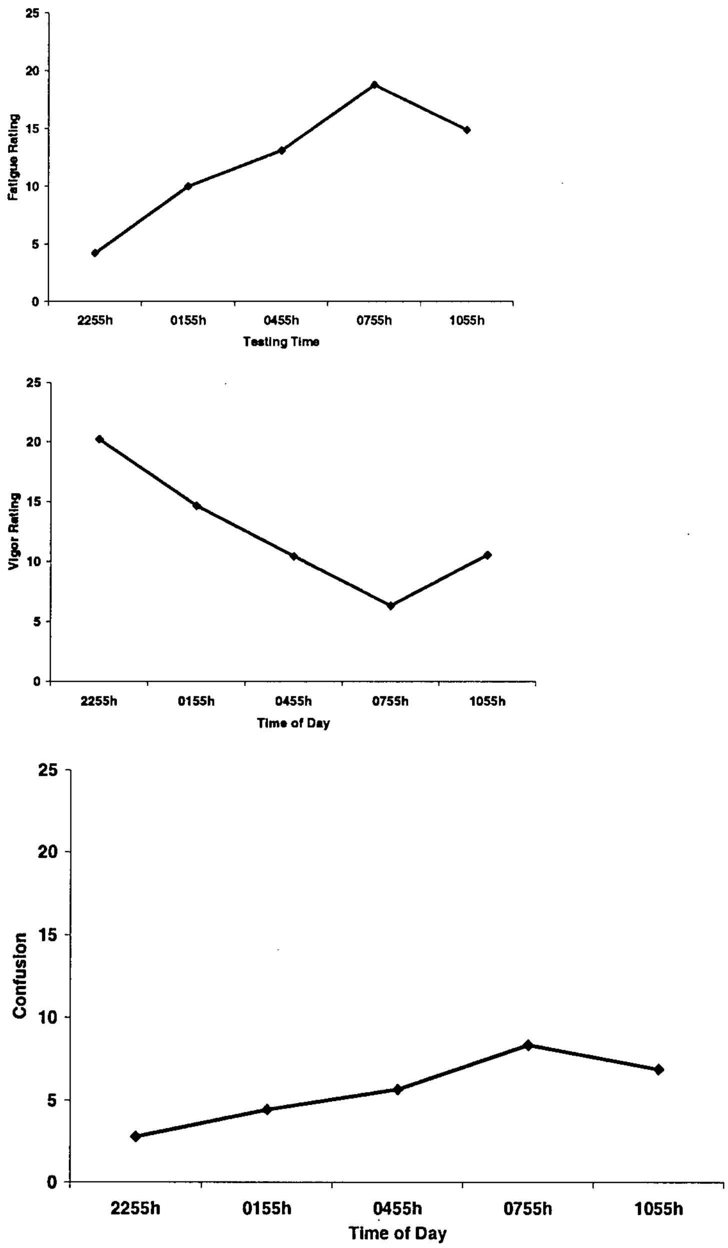

Low
High
8
7
6
5
Number Completed
2110h
0110h
0410h
0710h
1010h
Testing Time
Low
High
0.25
0.15
0.05
-0.05
Mean False Alarms
2110h
0110h
0410h
0710h
1010h
Testing Time
8
7.5
7
6.5
6
Reaction Tiime (seconds)
2210h
0110h
0410h
0710h
1010h
Testing Time

Figure 4.



25
20
15
10
5
0
Fatigue Rating
2255h
0155h
0455h
0755h
1055h
Testing Time
25
20
15
10
5
0
Vigor Rating
2255h
0155h
0455h
0755h
1055h
Time of Day
25
20
15
10
5
0
Confusion
2255h
0155h
0455h
0755h
1055h
Time of Day

Figure 5.

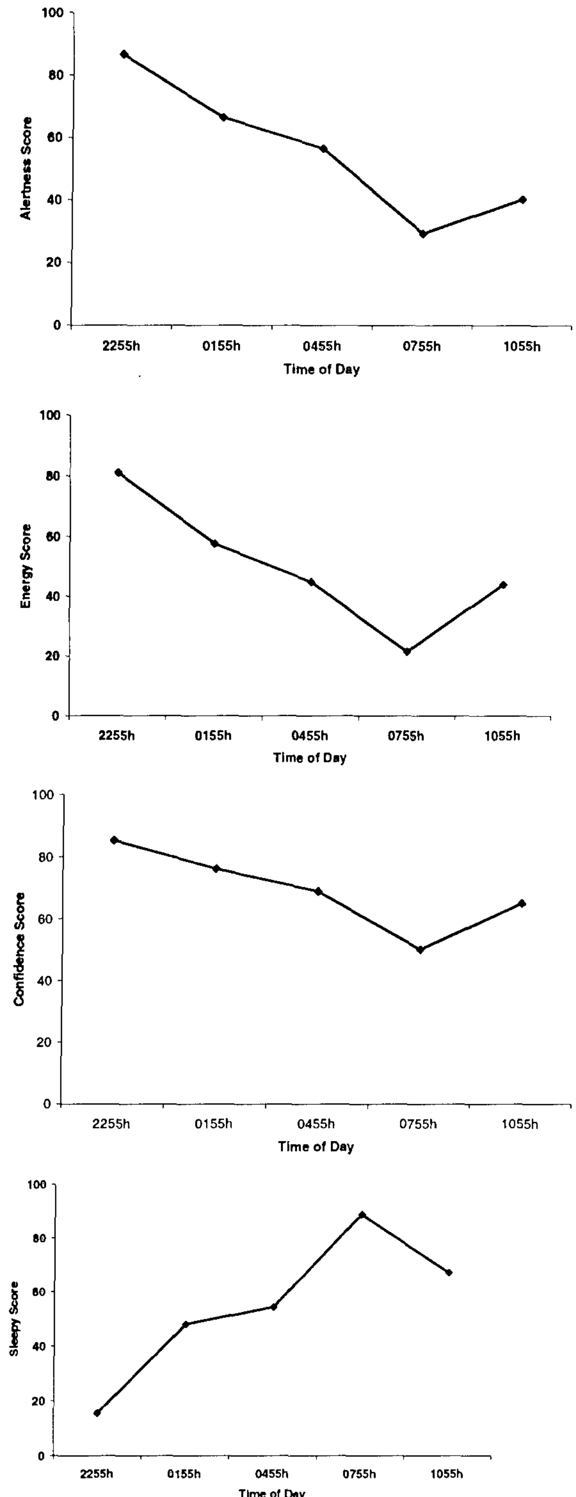

Alertness Score
Time of Day
2255h
0155h
0455h
0755h
1055h
Energy Score
Time of Day
Confidence Score
Time of Day
Sleepy Score
Time of Day

Figure 6.

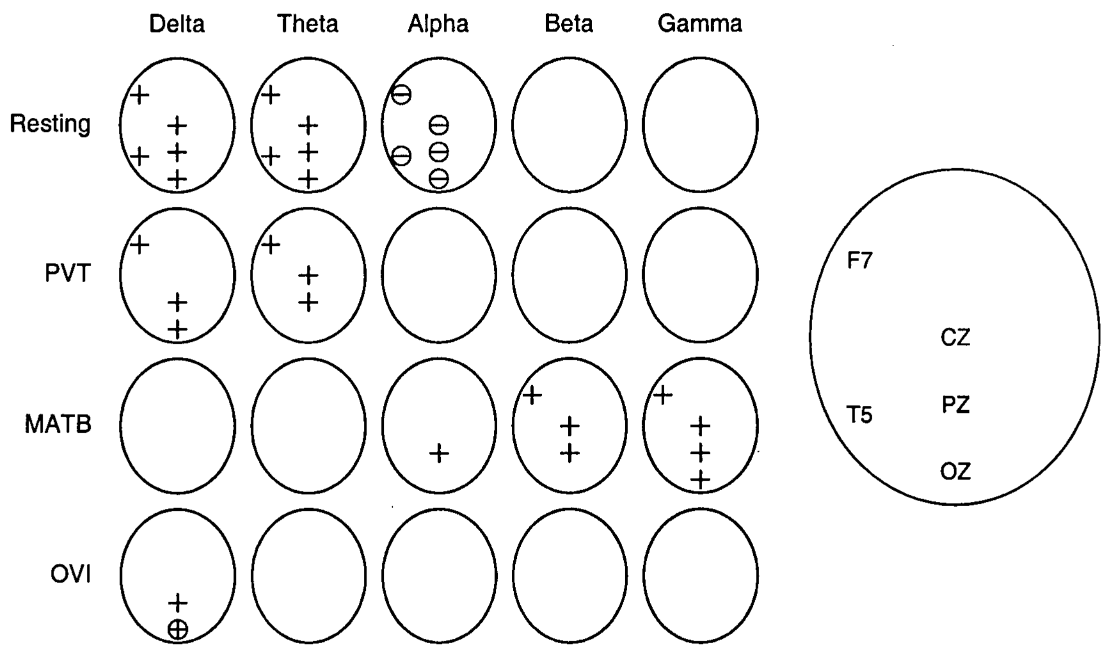

Delta
Theta
Alpha
Beta
Gamma
Resting
PVT
MATB
OVI
F7
CZ
T5
PZ
OZ

Figure 7.

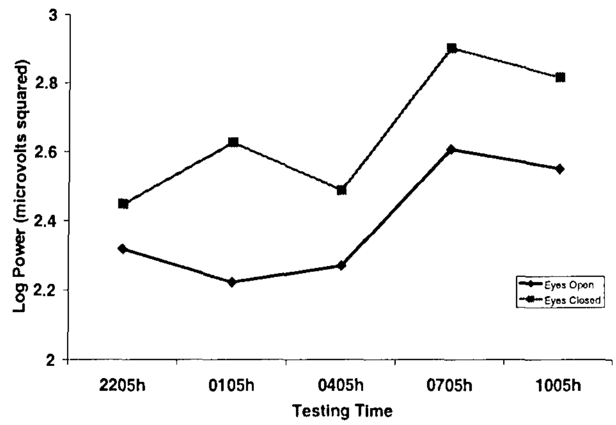


Figure 8.

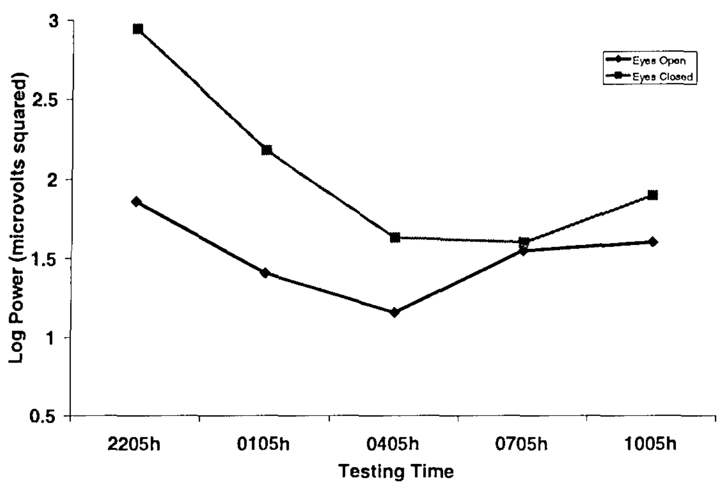


Figure 9.

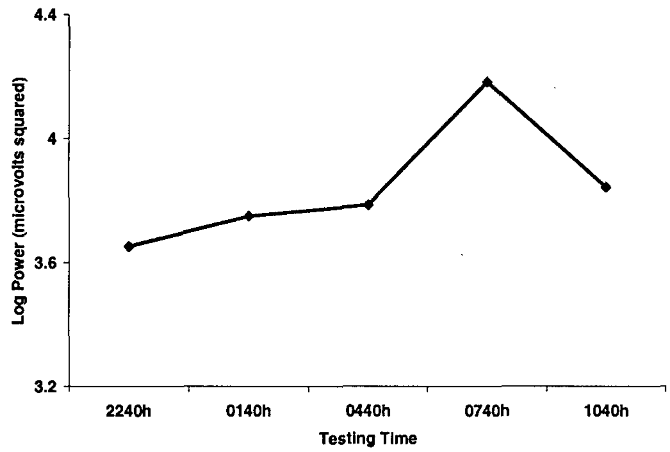


Figure 10.

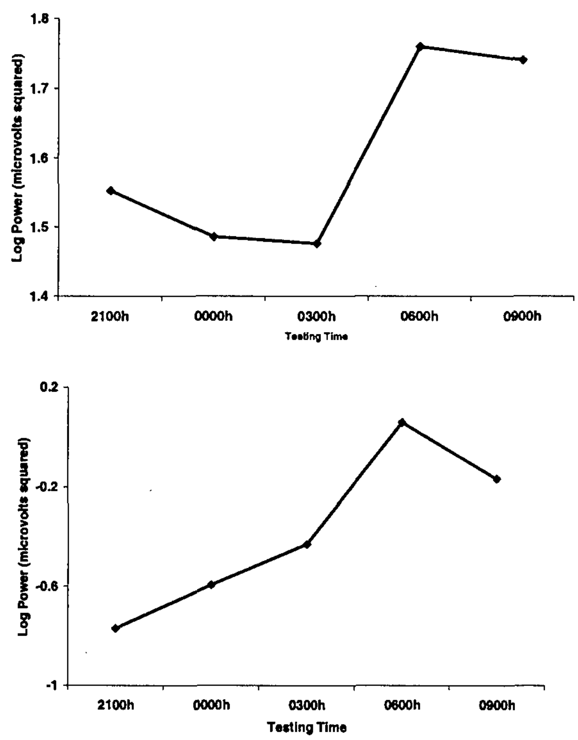


Figure 11.

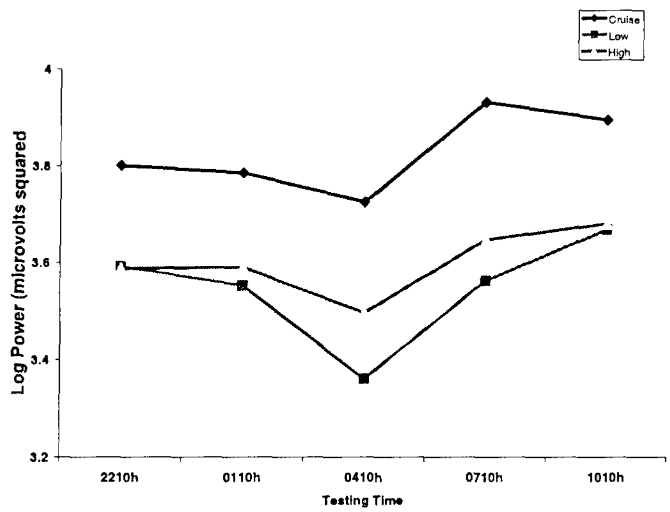


Figure 12.

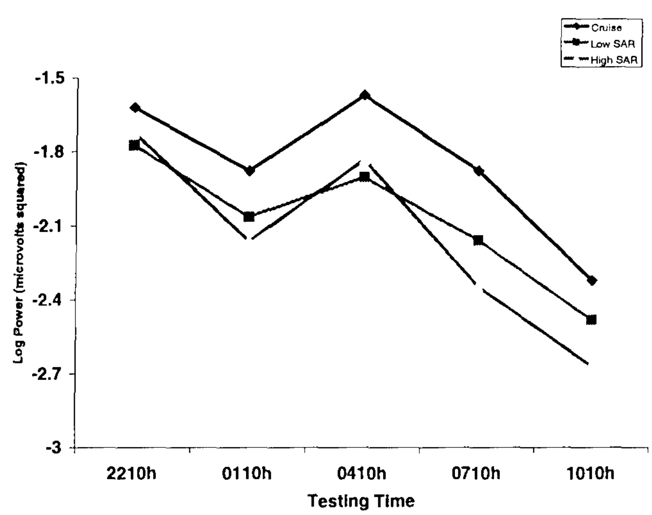


Figure 13.

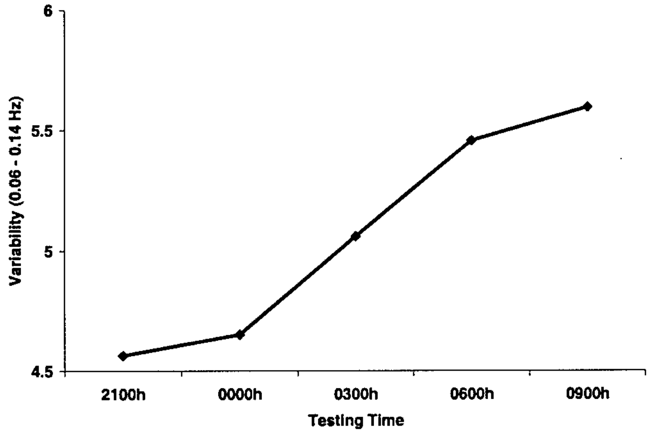

6
5.5
5
4.5
Variability (0.06 - 0.14 Hz)
2100h
0000h
0300h
0600h
0900h
Testing Time

Figure 14.